Soccer Broadcast Video Summarization

Khurram Soomro
Levi Smith
Christian Weigandt

Faculty Mentors:
Dr. Mubarak Shah
Dr. Niels Lobo

University of Central Florida
Center for Research in Computer Vision
crcv.ucf.edu
Shmuel Peleg’s Papers

- Creates compact summarization of events, moving temporally-distinct actions into the same time-space

- Use single stationary cameras

- Group activities do not get well-represented by the synopsis

- Disregards temporal ordering of actions in summary

- Not good for edited videos
Other Sports’ Summarization Papers

1. Summarize by detecting scene cuts/transitions and appearance of certain features
   - Not robust to different broadcasting styles

2. Summarize by detecting scenes which contain 'plays'
   - Not applicable to sports with continuous play such as soccer or hockey

3. Features include information such as camera angle or players on the field
   - Do not learn actions. Only learn to extract important frames.
## Action Occurrences

<table>
<thead>
<tr>
<th>Action</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arguing</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>5</td>
<td>4</td>
<td>9</td>
<td>35</td>
</tr>
<tr>
<td>Attempt At Goal</td>
<td>47</td>
<td>39</td>
<td>25</td>
<td>44</td>
<td>31</td>
<td>38</td>
<td>43</td>
<td>35</td>
<td>49</td>
<td>351</td>
</tr>
<tr>
<td>Card Red Yellow</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>Celebrating</td>
<td>5</td>
<td>20</td>
<td>10</td>
<td>13</td>
<td>12</td>
<td>9</td>
<td>7</td>
<td>13</td>
<td>10</td>
<td>99</td>
</tr>
<tr>
<td>Corner</td>
<td>13</td>
<td>7</td>
<td>11</td>
<td>11</td>
<td>7</td>
<td>11</td>
<td>15</td>
<td>10</td>
<td></td>
<td>97</td>
</tr>
<tr>
<td>Foul</td>
<td>36</td>
<td>20</td>
<td>21</td>
<td>29</td>
<td>59</td>
<td>17</td>
<td>42</td>
<td>28</td>
<td>36</td>
<td>288</td>
</tr>
<tr>
<td>Free Kick</td>
<td>16</td>
<td>9</td>
<td>12</td>
<td>17</td>
<td>25</td>
<td>9</td>
<td>21</td>
<td>9</td>
<td>21</td>
<td>139</td>
</tr>
<tr>
<td>Goal</td>
<td>11</td>
<td>36</td>
<td>21</td>
<td>24</td>
<td>26</td>
<td>12</td>
<td>16</td>
<td>9</td>
<td>15</td>
<td>170</td>
</tr>
<tr>
<td>Goal Kick</td>
<td>17</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>9</td>
<td>66</td>
</tr>
<tr>
<td>Injury</td>
<td>6</td>
<td>2</td>
<td>15</td>
<td>5</td>
<td>15</td>
<td>7</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>61</td>
</tr>
<tr>
<td>Offside</td>
<td>7</td>
<td>11</td>
<td>10</td>
<td>13</td>
<td>12</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>6</td>
<td>70</td>
</tr>
<tr>
<td>Penalty</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Substitution</td>
<td>4</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>42</td>
</tr>
<tr>
<td>Throw In</td>
<td>31</td>
<td>17</td>
<td>24</td>
<td>29</td>
<td>20</td>
<td>16</td>
<td>34</td>
<td>35</td>
<td>23</td>
<td>1675</td>
</tr>
</tbody>
</table>
Causality Matrix

\[ p(A|B) = \frac{|A \cap B|}{|B|} \]

- \(|A \cap B|\) is defined as the number of times where \(A\) is the first event that occurs within \(t\) seconds of \(B\).
- \(|B|\) is defined as the number of times that action \(B\) occurred.
Causality Matrix

The image shows a causality matrix with various action occurrences including Arguing, Attempt_At_Goal, Card_Red_Yellow, Celebrating, Corner, Foul, Free_Kick, Goal, Goal_Kick, Injury, Offside, Penalty, Substitution, and Throw_In. The matrix is color-coded with a gradient from 0.05 to 0.4, indicating the strength of the causal relationship between different events.
Each node will be the predicted probability using SVM. Use causality matrix as weights on edges.
Directed Acyclic Graph

- Use iterative approximation approach to find the path with max flow

- After determining labels, we will use graph clustering to find important segments of the game

- Use ground truth of highlights to learn significance priors for each event
Currently

- Calculated STIP and MBH features for each of the action clips
- Generating codebook for bag of words
- Look into other possible approaches such as HMM, CRF, etc.